
 W E S E R E (TM)

Release 3.1A John F. Collins, BioComputing Research Unit.
 Copyright (c) 1993-1998 University of Edinburgh, U.K.
 Distribution rights by Oxford Molecular Ltd

MPsrch_pp protein - protein database search, using Smith-Waterman algorithm
 Run on: Sat May 13 09:23:13 2000; MasPar time 3.41 Seconds
 Tabular output not generated. 194.564 Million cell updates/sec

Title: >US-09-331-631-28
 Description: (1-28) from US09331631.pep
 Perfect Score: 225
 Sequence: 1 LDPTRQQLCQMRCCQDEKDPKQOQCK 28

Scoring table: PAM 150
 Gap 15

Searched: 188963 seqs, 23686106 residues

Post-processing: Minimum Match 0%
 Listing first 45 summaries

Database: a:geneseq35
 1:geneseqp

Statistics: Mean 20.583; Variance 75.953; scale 0.271

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	225	100.0	28	1 W62841	Stenocarpus sinuatus a	1.69e-15
2	119	52.0	625	1 W62830	Macadamia integrifolia	2.93e-04
3	117	52.0	525	1 W62831	Theobroma cacao antimi	4.68e-04
4	117	52.0	566	1 R20181	Sequence encoded by 67	4.68e-04
5	113	50.2	666	1 W62828	Macadamia integrifolia	1.19e-03
6	111	49.3	666	1 W62829	Macadamia integrifolia	1.88e-03
7	104	46.2	805	1 R80561	Murine Ah receptor pro	5.79e-02
8	96	42.7	590	1 W62832	Gossypium hirsutum ant	1.61e-00
9	81	36.0	918	1 R12223	Human androgen recepto	1.61e-00
10	81	36.0	919	1 W14783	Androgen receptor.	1.61e-00
11	81	36.0	919	1 P93109	Human androgen recepto	1.61e-00
12	80	35.6	539	1 W33628	Yeast transcriptional	2.00e+00
13	76	33.8	1420	1 W81025	AIb1 (Amplified in bre	4.73e+00
14	74	32.9	351	1 R31599	Chicken nov protein.	7.25e+00
15	74	32.9	1313	1 W60213	Spinocerebellar ataxia	7.25e+00
16	73	32.4	760	1 W29450	Programmed cell death	6.96e+00
17	72	32.0	86	1 W95078	GST-HD fusion protein	1.11e+01
18	72	32.0	86	1 W95073	GST-HD fusion protein	1.11e+01
19	72	32.0	94	1 W95080	GST-HD fusion protein	1.11e+01
20	72	32.0	94	1 W95075	GST-HD fusion protein	1.11e+01
21	72	32.0	108	1 W95076	Amino acid sequence of	1.11e+01
22	72	32.0	108	1 W95071	Amino acid sequence of	1.11e+01
23	72	32.0	171	1 W99022	Human huntingtin prote	1.11e+01

24	72	32.0	371	1 W73369	Epitope tagged TBP pro	1.11e+01
25	72	32.0	731	1 W96312	Human small conductanc	1.11e+01
26	72	32.0	736	1 W63717	Human hSK3 protein.	1.11e+01
27	72	32.0	3144	1 R58777	Protein encoded by Hun	1.11e+01
28	72	32.0	3144	1 W44742	Human huntingtin prote	1.11e+01
29	72	32.0	3144	1 W09871	Human huntingtin.	1.11e+01
30	72	32.0	3144	1 W36887	Previously undescribed	1.11e+01
31	70	31.1	919	1 P90996	Human androgen recepto	1.69e+01
32	69	30.7	591	1 W48796	Drosophila melanogaster	2.08e+01
33	69	30.7	732	1 W63715	Rat rSK3 protein	2.08e+01
34	69	30.7	2441	1 W40058	Cellular transcription	2.08e+01
35	69	30.7	2441	1 R79054	CEB binding protein.	2.08e+01
36	68	30.2	737	1 W18317	Drosophila Deltex prot	2.56e+01
37	68	30.2	737	1 R76640	Deltex protein.	2.56e+01
38	68	30.2	737	1 R76639	Deltex protein.	2.56e+01
39	68	30.2	757	1 R73016	Fish protein gene tr	2.56e+01
40	68	30.2	1447	1 W81029	Murine pcip protein (o	2.56e+01
41	67	29.8	678	1 R42087	D. melanogaster dorsa	3.15e+01
42	67	29.8	914	1 W24800	Spinocerebellar ataxia	3.15e+01
43	67	29.8	1312	1 W33807	Human ataxin-2.	3.87e+01
44	66	29.3	902	1 R12224	Rat androgen receptor.	3.87e+01
45	66	29.3	1829	1 Y07242	Actin-filament binding	3.87e+01

ALIGNMENTS

RESULT 1
 ID W62841 standard; Protein: 28 AA.

AC W62841:
 DT 27-OCT-1998 (first entry)
 DE Stenocarpus sinuatus antimicrobial protein.
 KW antimicrobial protein; infestation; control.
 OS Stenocarpus sinuatus.
 PN W09827805-A1.
 PD 02-JUL-1998.
 PF 22-DEC-1997; AU0874.
 PR 20-DEC-1996; AU-004275.
 PA (REF-) COOP RES CENT TROPICAL PLANT PATHOLOGY.
 PI Bower NJ, Goulter KC, Green JL, Manners JM, Marcus JP;
 DR WPI: 98-377279/32.
 PT Novel antimicrobial protein from e.g. Macadamia integrifolia -
 useful for controlling microbial infestations of plants or mammals
 PS Claim 1; Page 66; 96pp; English.
 CC The sequence is that of an antimicrobial protein which can
 CC be used to control microbial infestations in plants and mammalian
 CC animals.
 SO Sequence 28 AA:

Query Match 100.0%; Score 225; DB 1; Length 28;
 Best Local Similarity 100.0%; Pred. No. 1.69e-15;

Matches 28; Conservative 0; Mismatches 0; Indels 0; Caps 0;

DB 1 LDPTRQQLCQMRCCQDEKDPKQOQCK 28
 QY 1 LDPTRQQLCQMRCCQDEKDPKQOQCK 28

RESULT 2

ID W62830 standard; Protein: 625 AA.

AC W62830;
 DT 27-OCT-1998 (first entry)
 DE Macadamia integrifolia antimicrobial protein.
 KW antimicrobial protein; infestation; control.
 OS Macadamia integrifolia.

FH key Location/Qualifiers

FT Peptide 1..28

FT Protein /note= "signal peptide"

FT /note= "666"

FT /note= "mature protein"

PN W09827805-A1.

PD 22-DEC-1997; AU0874.

PR 20-DEC-1996; AU-004275.

PA (RETR-). COOP RES CENT TROPICAL PLANT PATHOLOGY.
PI Bower NI, Goulter KC, Green JL, Manners JM, Marcus JP;
DR WPI: 98-377279/32.
DR N-PSDB: V42316
PT Novel anti-microbial protein from e.g. *Macedamia integrifolia* -
PT useful for controlling microbial infestations of plants or mammals
PS Claim 1; Page 43-45; 9pp; English.
CC The sequence is that of an antimicrobial protein which can
CC be used to control microbial infestations in plants and mammalian
CC animals.
SQ Sequence 625 AA;

Query Match	52.9%;	Score 119;	DB 1;	Length 625;
Best Local Similarity	65.4%;	Pred. No. 2.93e-04;		
Matches	17;	Conservative	2;	Mismatches 7; Indels 0; Gaps

```
Db      36 DPQTCCQCQRCRCRQGESDPQQQYC 61
      || : | | | | : | | | | |
Qy      2 DPIRQQQLCOMRCQQQEQKDPQQQQC 27
```

RESULT	3	
ID	M62831	standard; Protein; 525 AA.
AC	M62831.	
DT	27-OCT-1998	(first entry)
DE	Theobroma cacao antimicrobial protein.	
KW	antimicrobial protein; infestation; control.	
OS	Theobroma cacao.	
PN	WC9827805-A1.	
PD	02-JUL-1998.	
PE	22-DEC-1997.	
PR	20-DEC-1996.	AU-004275.
FA	(RETR-) COOP RES CENT	TROPICAL PLANT PATHOLOGY.
P1	Bower NI, Goulier KC, Green JL, Manners JM, Marcus JP;	
DR	WPI; 98-3177219/32.	
PT	Novel anti-microbial protein from e.g. <i>Macadamia integrifolia</i> -	
PT	useful for controlling microbial infestations of plants or mammals	
PS	Claim 1; Page 47-49; 96pp; English.	
CC	The sequence is that of an antimicrobial protein which can	
CC	be used to control microbial infestations in plants and mammalian	
CC	animals.	
CC	Sequence	525 AA;
10		

	Query Match	52.0%	Score 117;	DB 1;	length 525;
	Best Local Similarity	60.9%	Pred No.	4,	68-04;
	Matches	14;	Conservative	3;	Mismatches 6;
				Indels	0;
				Gaps	0;
Dd	83	RQYQQCCGRCQEQDQGAGREQQQC	105		
Oy	5	RQQLCQMRCCQQLKDKPRQHQDQC	27		

RESULT 4
ID R20181 standard; Protein; 566 AA.
AC R20181:
DT 16-APR-1992 (first entry)
DE Sequence encoded by 67 kD T. cacao protein cDNA.
KW Cocoa; flavour; vicillin; seed storage protein.
OS Theobroma cacao.
PN MO9119801-A.
PD 26-DEC-1991.
PE 07-JUN-1991: G00914.
PR 11-JUN-1990: GB-013016.
RA (MASC) MARS UK LTD.
PI Spencer ME, Hodge R, Deakin EA, Ashton S;
DR WPI: 92-024418/03.
PT N-PSDB: Q20377.
PT beans and produced in large quantities using yeast and bacterial
PT expression vectors
PS Claim 4; Fig 2; 59pp; English.
CC The inventors claim a 67 kD and 31 kD T. cacao protein, and
fragments, and encoding DNAs. The 47 kD and 31 kD proteins are

CC deduced from the 67 kD precursor. T. cacao protein cDNA was
CC detected in a cDNA library prepared from immature cocoa beans RNA
CC using a probe based on the AA sequence of a CNBR peptide common to
CC the 47 kD and 31 kD polypeptides. Homology searches revealed close
CC homologues between the 67 kD polypeptide and the vicilins, which are
CC seed storage proteins.
CC Sequence 566 AA:

Query Match	52.0%;	Score 117;	DB 1;	Length 566;
Best Local Similarity	60.9%;	Pred. No. 4.68e-04;		
Matches	14;	Conservative	3;	Mismatches 6; Indels 0; Gaps 0;

```
Db      83 RQYQCCGRCQEQQQGGREDDC 105
        ||| || |||:| |:|||
Qy      5  RQQQLCQMRCCQQQEKDPRQQQC 27
```

RESULT	5	Mac2828 standard; Protein; 666 AA.
ID	W62828	
AC	W62828;	
DT	27-OCT-1998	(first entry)
DE	Macadamia integrifolia antimicrobial protein.	
KW	antimicrobial protein; infestation; control.	
OS	Macadamia integrifolia.	
FH	key	Location/Qualifiers
FT	peptide	1..28
FT		/note="signal peptide"
FT	protein	29..666
FT		/note="mature protein"
PN	W09827805-A1.	
PD	02-JUL-1998.	
PF	22-DEC-1997; A00874.	
PR	20-DEC-1996; AU-004275.	
PA	(RETR-) COOP RES CENT TROPICAL PLANT PATHOLOGY.	
PI	Bower N.I., Goulter KC, Green JL, Manners JM, Marcus JP;	
DR	W62828; W42310.	
N	N-PSDB; W42310.	
FT	Novel anti-microbial protein from e.g. Macadamia integrifolia -	
PT	useful for controlling microbial infestations of plants or mammals	
PS	Claim 1; Page 34-36; 96pp; English.	
CC	The sequence is that of an antimicrobial protein which can	
CC	be used to control microbial infestations in plants and mammalian	
CC	animals.	
Q	Sequence	666 AA;

		50.2%	Score 113;	DB 1;	Length 666;
Query Match					
Best Local Similarity		61.5%;	Pred. No. 1, 19e-03;		
Matches	16;	Conservative	2;	Mismatches 8;	Indels 0; Gaps 0.
Ddb	77 DPOTECQCCQRRCRQDSGPRQQGYC	102			
	:				
OY	2 DPIRQQQLCOMRCQOQEKDPPROQOC	27			

RESULT	6
ID	W62829 standard; Protein; 666 AA.
AC	W62829;
DT	27-OCT-1998 (first entry)
DE	Macadamia integrifolia antimicrobial protein.
KW	antimicrobial protein; infestation; control.
OS	Macadamia integrifolia.
FH	Key location/qualifiers
FT	Peptide 1..28
FT	/note="signal peptide"
FT	29..666
FT	/note="mature protein"
PN	W09827805-A1.
PD	02-JUL-1998.
PF	22-DEC-1997: AU0874.
PR	20-DEC-1996: AU-004275.
PA	(RETR-) COOP RES CENT TROPICAL PLANT PATHOLOGY.
PI	Bower NJ, Goulter KC, Green JL, Mannens JW, Marcus JP.
WI	WPI: 98-371279/32.

.J.S. Homo sapiens.
.L.N. M09711170-A1.
.PD 27-MAR-1997.
.PF 20-SEP-1996; U15081.
.PR 20-SEP-1995; US-004018.
.PA (WORC-) WORCESTER FOUND BIOMEDICAL RES.
.P1 Zamecnik PA;
.DR WPI: 97-202879/18.
.DR N-PSDB: T63407.
.PT Oligonucleotide(s) antisense to human androgen receptor and acidic
.PT Fcγ genes - used to inhibit gene expression, for the treatment of
.PT benign prostatic hyperplasia
.PS Disclosure; Page 22-28; 51pp; English.
.CC Human androgen receptor (W14783) binds testosterone and, acting
at the transcriptional level, regulates the growth of normal
prostatic cells. Antisense oligonucleotides (see also T63200,
T63404-05) based on an androgen receptor cDNA clone (see also
T63407) can be used to prevent androgen receptor gene expression,
thereby inhibiting the growth or survival of prostatic cells for
the treatment of benign prostatic hyperplasia and prostate cancer.
Sequence 919 AA;

Query Match	36.08;	Score 81;	DB 1;	Length 919;
Best Local Similarity	63.68;	Pred. No. 1.61e+00;		
Matches	14;	Conservative	2;	Mismatches 6;
				Indels 0;
				Gaps 0

```
Db 66 QQQQQQQQQQQQQQETSPRQQQ 87
   :||| | : ||| | |||||
Qy 5 RQQQLCQMRCCQQEKDPRQQQ 26
```

RESULT	11
ID	P93109 standard; protein; 919 AA.

DT 19-MAR-1990 (first entry)
DE Human androgen receptor.
KW Human androgen receptor; ployclonal antibody; cancer.
OS Homo sapiens.
PN W0809791.A.
PD 19-OCT-1989.
PF 13-APR-1989; U01548.
PR 14-APR-1988; US-182646.
PA (UNC-) University of North Carolina.
PI French FS, Wilson EM, Joseph DR, Lubahn DB;
DR WPI: 89-324206/44.
DR N-PSDB: N911772.
PT DNA encoding androgen receptor protein - useful for transforming
PT eukaryotic hosts for protein expression and subsequent antibody prodn.
PS Disclosure: Fig. 4; 41pp; English.
CC Androgen receptor protein (AR) is used to produce mono- or poly-clonal
CC antibodies. These are used for the detection and quantification of AR in
CC the presence of endogenous androgen, as androgen will not interfere with
CC binding. They may be used in assays to determine and quantify cellular
CC distribution of AR in tumour tissue, and are esp. useful for evaluating
CC prostate cancers to determine responsiveness to androgen withdrawal
CC therapy.
CQ Sequence 919 AA;

Query Match	36.08;	Score 81;	DB 1;	Length 919;
Best Local Similarity	63.68;	Pred. No. 1.61e+00;		
Matches	14;	Conservative 2;	Mismatches 6;	Indels 0;
			Gaps	0;

```

Db      66 QQQQQQQQQQQQQQETSPRQQQ 87
      :||| | : ||| | |||||
Qy      5 RQQQLCQMRCCQQQEKDPQQQQ 26

```

RESULT 12
ID W33628 standard; Protein: 539 AA.
AC W33628;
DT 06-JUL-1998 (first entry)
DE yeast transcriptional activator factor TAF-68 protein.
KW TAF-68; transcriptional activator factor; transactivating factor.

KW TATA-box binding protein-associated factor:yeast:
 KM transcription, viability, antifungal, fungicide, infection:
 KN therapy.
 OS Saccharomyces cerevisiae.
 PN M09749828-A1.
 PD 31-DEC-1997.
 PE 26-JUN-1997: U11144.
 PR 26-JUN-1996: US-673234.
 PA (UYMA-) UNIV MASSACHUSETTS.
 PI Green MR, Reese JC:
 DR WPL: 98-077189/07.
 PI N-PSDB: V04771.
 PT 68 - nucleic acids encoding TATA-box binding protein factors 47 and
 PT yeast nucleic acids encoding TATA-box binding protein factors 47 and
 PT useful for developing anti-fungal agents for immunocompromised
 PT patients
 PS Claim 9: Page 54-57, 93pp: English.
 CC This polypeptide comprises a 68 kDa yeast transactivating factor,
 CC TAR-66, that is associated with the TATA-box binding protein.
 CC TAR-66 can be isolated from *Saccharomyces cerevisiae* cells by
 CC virtue of its affinity to fungal or human TARA-box binding proteins
 CC using chromatographic procedures. It can also be recombinantly
 CC produced in host cells utilising vectors carrying a TAR-47 nucleic
 CC acid (see V04771). The yeast TAR complex includes 9 polypeptides
 CC of 180, 145 (see W33632), 116, 90 (see W33634), 68, 62, 51-54, 47
 CC (see W33627) and 30 kDa. At least some of these are required for
 CC transcription in vitro. TARs are also essential for viability.
 CC TAR-66 can be used in the discovery, design and development of
 CC antifungal agents, e.g. for treatment of *Candida albicans* infections
 CC (candidiasis). It can also be used to raise diagnostically useful
 CC anti-TAR antibodies.
 SO Sequence 539 AA:

Query Match	35.6%;	Score 80;	DB 1;	Length 539;
Best Local Similarity	44.0%;	Pred. No. 2.00e+00;		
Matches	11;	Conservative	8;	Mismatches 6;
				Indels 0;
				Gaps 0;

```

Db      241 ESTQQQRVVQQQRVVQQQQQQQQQQQQQ 265
      :::::| | | | : : | | |
QY      2 DPIRQQQLCQMRCQQQEKDPRQQQ 26

```

ID	WT81025	standard;	Protein:	1420	AA.
RESULT	13				
DE	12-MAY-1999	(first entry)			
DE	AlBI (Amplified in breast cancer-1) protein.				
DE	AlBI; amplified in breast cancer; cancer; steroid; receptor;				
KW	coactivator; SCR; estrogen; ER; estrogen dependent transcription				
KW	breast cancer; lung cancer; colon cancer; prostate cancer;				
KW	melanoma.				
OS	Homo sapiens.				
FT	Key				
FT	Location/Qualifiers				
FT	16..88				
FT	/label= Basic_helix-loop_helix_domain				
FT	116..111				
FT	/label= PAS_domain				
FT	263..302				
FT	/label= PAS_domain				
FT	396..715				
FT	/label= Serine/Threonine_rich_region				
FT	767..888				
FT	/label= Serine/Threonine_rich_region				
FT	1129..1244				
FT	/label= Glutamine_rich_region				
FT	1244..1272				
FT	/label= Polyglutamine_tract				
PN	MO9857982-A2.				
PN	23-DEC-1998.				
PF	17-JUN-1998; U12689.				
PR	17-JUN-1997; US-049728.				
PA	(USGO) US GOVERNMENT.				
PI	Meltzer P, Trent JM;				
PR	WPI: 99-080946/07.				

DR N-PSD: V99915.
PT New isolated steroid receptor co-activator, AIB1 - used to develop
PT products for the diagnosis and treatment of steroid-responsive
PT tumours, e.g. breast, lung, prostate or colon cancers or melanomas
PS Claim 3; Page 31-33; 57pp; English.
CC The AIB1 protein is a member of the steroid receptor coactivator-1
CC (SRC-1) family of nuclear receptor co-activators that interact with
CC estrogen receptors (ER) to enhance ER-dependent transcription.
CC The AIB1 gene is amplified and over-expressed in certain cancers
CC in particular breast cancer and steroid hormone responsive cancers.
CC The AIB1 polypeptide can be used to identify compounds which inhibit
CC ER-dependent transcription. Increased expression of the AIB1 gene
CC indicates aberrantly proliferating cells, thus detection of
CC increased expression of the AIB1 gene or an increase in the number
CC of copies of the AIB1 gene can be used to diagnose cancer or a
CC predisposition towards developing cancer. Compounds which inhibit
CC expression of AIB1 or compounds which inhibit interaction of AIB1
CC with steroid receptors or nuclear co-factors can be used for
CC reducing the proliferation of cancer cells.
SQ Sequence 1420 Aa.

Query Match	33.88;	Score 76;	DB 1;	Length 1420;
Best Local Similarity	47.88;	Pred. No. 4.73e+00;		
Matches	11;	Conservative	6;	Mismatches 6;
				Indels 0;
				Gaps 0;

```
Db 1234 FRQQRVAMMMQZZZZZZZZZZ 1256
      :|||: | |||: : ||||
QY 4 IRQQQLCQMRCCQQQEKDPRQQQ 26
```

RESULT	14
ID	R31599 standard; Protein: 351 AA.
AC	R31599;
NT	24-MAY-1993 (first entry)
DE	Chicken nov protein.
KW	avian nephroblastoma; avian myeloblastoma virus; IGF binding site;
KW	insulin-like growth factor; Wilm's tumour.
OS	Gallus domesticus.
FX	
FT	Key
FT	Location/Qualifiers
FT	peptide
FT	1..24
FT	/label="signal_peptide
FT	/note="only hydrophobic region of protein"
FT	56..63
FT	/label="IGF-binding_site_motif
FT	/note="corresponds to GCGCCXC consensus"
FT	
PN	W09300430-A.
PD	07-JAN-1993.
PF	25-JUN-1992; F00589.
PR	25-JUN-1991; FR-007807.
PI	(CMRS) CENT NAT RECH SCL.
PI	Marlunerie C, Perdal B;
DR	WPI: 93-03637/04.
DR	N-PSDB: Q36031.
PT	Nucleotide sequences hybridising to regions of chicken nov gene -
PT	useful as probes for detecting complementary sequences to
PT	evaluate development and/or differentiation of tumours
PS	Claim.1; Fig 1; 67bp; French.
CC	This amino acid sequence was deduced from the nucleotide sequence
CC	of a chicken nov gene clone isolated from a gene bank prepared from
CC	chicken embryonic fibroblasts screened with a tumour-derived probe.
CC	The only hydrophobic region occurs within the putative signal
CC	peptide suggesting that the protein is secreted. The protein also
CC	contains the consensus motif of proteins which bind to insulin-like
CC	growth factors. It is known that the human IGFRI gene is
CC	overexpressed in some Wilm's tumours and a similar deregulation of
CC	IGFRI expression could be involved in nephroblastoma development.
CC	The deduced nov protein sequence contains 39 (non-clustered)
CC	cysteine residues.
CC	Sequence 351 AA.

Query Match	32.9%	Score 74;	DB 1;	Length 351;
Best Local Similarity	29.6%;	Pred. NO. 7.25e+00;		
Matches	8;	Conservative	7;	Mismatches 11; Indels 1; Gaps 1;

```

Db      232 EMVKQTRLCMRRCENEPEPSDKKGGKC 258
          : : : | | | : : |
QY      2  DPIRQQQLCMR-CQQQEKDPQQQQC 27

```

RESULT	15	
ID	M60213	standard; Protein: 1313 AA.
AC	M60213:	
DT	02-OCT-1998	(first entry)
DE	Spinocerebellar ataxia type 2 (SCA2) disease associated protein.	
KW	Spinocerebellar ataxia type 2; SCA2; gene therapy; antisense therapy;	
RV	CAG repeat; neurodegenerative disease.	
OS	Homo sapiens.	
EH	Key	Location/Qualifiers
FT	Region	166..188
FT		/note="Glutamine rich region; this region is
FT		increased in SCA2 patients"
PD	W09818920-A1.	
PN	07-MAY-1998.	
PT	30-OCT-1997:	J03946.
PR	30-OCT-1996:	JP-304059.
PA	(SRLS-) SRL INC.	
PI	Sampei K, Tsuji S;	
DR	Wpi: 98-272215/24.	
PT	N-PSDB: V30270.	
FT	Nucleic acid fragments associated with spinocerebellar ataxia type 2	
FT	- contain increased number of CAG repeat region compared to normal	
FT	gene	
PS	Claim 1; Pages 13-22; 38pp;	Japanese.
CC	This is the protein sequence of a gene causative of spinocerebellar	
CC	ataxia type 2 (SCA2), a neurodegenerative disease. The gene associated	
CC	with SCA2, has a tri-nucleotide (CAG) repeat region which in the	
CC	expression product produces a polyglutamine sequence from Gln-156 to	
CC	Gln-188. In the normal gene there are 15-25 CAG repeats but in SCA2	
CC	patients this number is increased to 35-100. Peptides encoded by nucleic	
CC	acid fragments (DNA or RNA) containing sequences from the SCA2 associated	
CC	gene, antibodies recognizing the peptides and antisense nucleic acids	
CC	hybridising with the nucleic acid fragments can be used for the	
CC	investigation and diagnosis of SCA2. They can also be used for the	
CC	treatment of SCA2 by antisense therapy.	
CC	Sequence 1313 AA:	

Query Match	32.9%;	Score 74;	DB 1;	Length 1313;
Best Local Similarity	50.0%;	Pred. No. 7.25e+00;		
Matches	13;	Conservative	5;	Mismatches 8; Indels 0; Gaps 0;

[illegible]

Search completed: Sat May 13 09:23:21 2000
Job time : 8 secs.

